JOURNAL OF BACTERIOLOGY, June 1990, p. 3335–3345
0021-9193/90/063335-11$02.00/0

Vol. 172, No. 6

Genes for Two Herbicide-Inducible Cytochromes P-450 from *Streptomyces griseolus*†

CHARLES A. OMER,[1]* REIJER LENSTRA,[1]‡ PATRICIA J. LITLE,[1] CAROLINE DEAN,[2]§ JAMES M. TEPPERMAN,[2]‖ KENNETH J. LETO,[3] JAMES A. ROMESSER,[1] AND DANIEL P. O'KEEFE[1]

Central Research and Development Department, Du Pont Experimental Station, E. I. Du Pont Co., Wilmington, Delaware 19880[1]; Agricultural Products Department, Stine-Haskell Laboratories, E. I. Du Pont Co., Newark, Delaware 19711[3]; and DNA Plant Technology, Oakland, California 94608[2]

Received 17 November 1989/Accepted 23 March 1990

Streptomyces griseolus **ATCC 11796 contains two inducible, herbicide-metabolizing cytochromes P-450 previously designated P-450$_{SU1}$ and P-450$_{SU2}$ (P-450CVA1 and P-450CVB1, respectively, using nomenclature of Nebert et al. [D. W. Nebert, M. Adesnik, M. J. Coon, R. W. Estabrook, F. J. Gonzalez, F. P. Guengerich, I. C. Gunsalus, E. F. Johnson, B. Kemper, W. Levin, I. R. Phillips, R. Sato, and M. R. Waterman, DNA 6:1–11, 1987]). Using antibodies directed against cytochrome P-450$_{SU1}$, its N-terminal amino acid sequence, and amino acid composition, we cloned the *suaC* gene encoding cytochrome P-450$_{SU1}$. Similar information about the cytochrome P-450$_{SU2}$ protein confirmed that a gene cloned by cross-hybridization to the *suaC* gene was the *subC* gene encoding cytochrome P-450$_{SU2}$. The *suaC* and *subC* genes were expressed in *Escherichia coli*, DNA for both genes was sequenced, and the deduced amino acid sequences were compared with that of the well-characterized cytochrome P-450$_{CAM}$ from *Pseudomonas putida*. Both cytochromes P-450$_{SU1}$ and P-450$_{SU2}$ contain several regions of strong similarity with the amino acid sequence of P-450$_{CAM}$, primarily in regions of the protein responsible for attachment and coordination of the heme prosthetic group.**

Streptomyces griseolus ATCC 11796 is capable of metabolizing a number of sulfonylurea herbicides to compounds that often exhibit reduced phytotoxicity (23, 24, 30). It does so via two sulfonylurea-inducible cytochrome P-450 monooxygenases designated cytochrome P-450$_{SU1}$ and cytochrome P-450$_{SU2}$ (24). Partial characterization and reconstitution studies (23, 24) suggest that the two inducible P-450 monooxygenase systems of *S. griseolus* resemble the three-component cytochrome P-450$_{CAM}$ camphor oxidation system of *Pseudomonas putida* ATCC 17453 (12). In the P-450$_{CAM}$ system, the genes for the three components, putidaredoxin reductase, putidaredoxin (iron-sulfur protein), and cytochrome P-450$_{CAM}$, have been designated *camA*, *camB*, and *camC*, respectively (15, 28). In line with this genetic nomenclature and anticipating the genetic characterization of all the components of the *S. griseolus* systems, we propose to name the gene for cytochrome P-450$_{SU1}$ *suaC* and the gene for cytochrome P-450$_{SU2}$ *subC*.

We are interested in analyzing the genes for these sulfonylurea-metabolizing systems and how they are regulated to understand how a soil organism responds to and metabolizes chemicals foreign to its environment. Additionally, we would like to introduce these genes into plants to enable them to metabolize sulfonylureas to less phytotoxic compounds. In this study, we used the characteristics of the cytochrome P-450$_{SU1}$ and P-450$_{SU2}$ apoproteins (amino acid composition, NH_2-terminal sequence, and antigenicity) and their relatedness to one another to identify and clone the genes *suaC* (P-450$_{SU1}$) and *subC* (P-450$_{SU2}$). We also sequenced the DNA encoding these two proteins, expressed them in *Escherichia coli*, and compared the deduced amino acid sequences of cytochromes P-450$_{SU1}$ and P-450$_{SU2}$ with that of the best-studied bacterial cytochrome P-450, that of P-450$_{CAM}$ in *P. putida*.

MATERIALS AND METHODS

Materials. General chemicals were obtained from a number of standard sources including Sigma Chemical Co. (St. Louis, Mo.) and International Biotechnologies, Inc. (New Haven, Conn.). Restriction enzymes and DNA modification enzymes were obtained from New England Bio-Labs, Inc. (Beverly, Mass.), Bethesda Research Laboratories, Inc. (Gaithersburg, Md.), Boehringer Mannheim Biochemicals (Indianapolis, Ind.), and U.S. Biochemical Corp. (Cleveland, Ohio). [α-^{32}P]dCTP (800 or 3,000 Ci/mmol), [α-^{32}P]dATP (800 Ci/mmol), [α-^{35}S]dATP (1,350 Ci/mmol), and [γ-^{32}P]ATP (3,000 Ci/mmol) were obtained from Dupont, NEN Research Products (Boston, Mass.). ^{125}I-labeled protein A (>30 mCi/mg) was obtained from ICN Radiochemicals (Irvine, Calif.). Sulfonylurea compounds were obtained from Du Pont Co. (Wilmington, Del.). Bacteriological medium was obtained from Difco Laboratories (Detroit, Mich.). Ampicillin, kanamycin, and tetracycline were obtained from Sigma. Reagents for immunodetection of proteins (Western blots) were obtained from Bio-Rad Laboratories (Richmond, Calif.).

Purification, preparation of antisera, N-terminal amino acid analysis, and amino acid composition of cytochromes P-450$_{SU1}$ and P-450$_{SU2}$. Cytochrome P-450$_{SU1}$ was purified as previously described (24), and anti-cytochrome P-450$_{SU1}$ rabbit polyclonal antibodies were prepared from protein electroeluted from polyacrylamide gels (23). Cytochrome P-450$_{SU2}$ was purified by a procedure identical to that used for P-450$_{SU1}$ (24) with the exception that the retention time on anion-exchange high-pressure liquid chromatography was

* Corresponding author.
† Du Pont Experimental Station Contribution 5241.
‡ Present address: Glaxo Institute for Molecular Biology S.A., Geneva 24, Switzerland.
§ Present address: John Innes Institute, Norwich NR4 7UH, England.
‖ Present address: U.S. Department of Agriculture/Plant Gene Expression Center Laboratories, Albany, CA 94710.

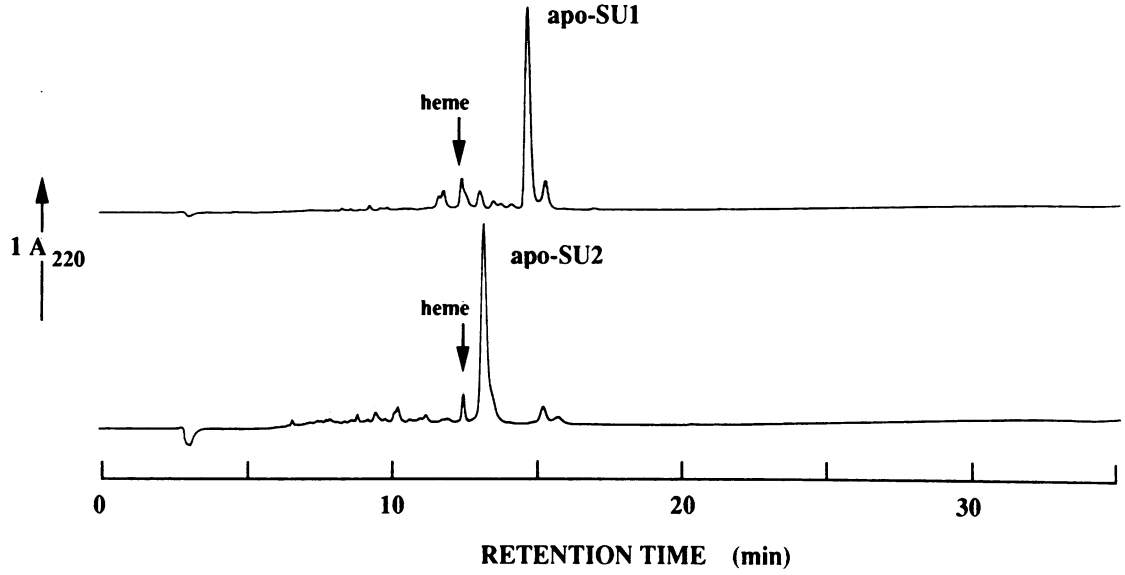

FIG. 1. Reverse-phase high-pressure liquid chromatography of purified cytochromes P-450$_{SU1}$ and P-450$_{SU2}$. About 2 nmol of each protein was injected. For each sample, apoprotein was collected while the detector indicated $A_{220} \geq 0.2$.

~26 min, compared with ~35 min for P-450$_{SU1}$. A typical preparation from 200 ml of cell paste containing 400 nmol of total cytochrome P-450 yielded ~40 nmol of cytochrome P-450$_{SU1}$ and ~10 nmol of cytochrome P-450$_{SU2}$. Denaturing polyacrylamide gel electrophoresis (M_r = 46,500 for P-450$_{SU1}$, and M_r = 46,000 for P-450$_{SU2}$), and reverse-phase high-pressure liquid chromatography (Fig. 1) were used to judge these proteins essentially homogeneous. Purified native cytochromes prepared in this manner exhibited room temperature A_{418}/A_{280} values of 1.72, and 1.78 for P-450$_{SU1}$ and P-450$_{SU2}$, respectively.

To prepare polyclonal antisera to cytochrome P-450$_{SU2}$, we electrophoresed on polyacrylamide gels the 30 to 60% ammonium sulfate fraction of protein extracts of *S. griseolus* PH2042 cells induced by chlorimuron ethyl plus 10 mM aniline. The approximately 44-kilodalton (kDa) band corresponding to cytochrome P-450$_{SU2}$ was cut out of the gel. Polyclonal antiserum directed against cytochrome P-450$_{SU2}$ was obtained after subcutaneous injection of powdered polyacrylamide gel slices containing cytochrome P-450$_{SU2}$ into New Zealand White rabbits. Antiserum raised against this 44-kDa protein bound to chromatographically purified cytochrome P-450$_{SU2}$ but did not cross-react with purified cytochrome P-450$_{SU1}$ (see Fig. 6) or other *S. griseolus* proteins in the 40- to 50-kDa range.

Reverse-phase chromatography and preparation of apoproteins was as described previously (22), using a Vydac C-4 column (type 214TP54, 4.5 by 250 mm) and utilizing a combination of linear gradients (flow = 1.0 ml/min) starting with 5% solvent B in solvent A, progressing from 5 to 40% solvent B between 1 to 5 min and from 40 to 100% solvent B between 5 and 30 min. Solvent A was 0.1% trifluoracetic acid in water; solvent B was 0.1% trifluoracetic acid in acetonitrile. Preparations of the apoproteins from this step were subjected to reduction with dithiothreitol and alkylation with iodoacetic acid to make the carboxymethylcysteine derivative (9) and, after rechromatography on the reverse-phase column, were used to determine amino acid compositions (1) and NH_2-terminal sequences (17).

Bacterial strains and plasmids. *E. coli* MC1061 (5) was used as a plasmid-cloning host; *E. coli* Y1088, Y1089, and Y1090 (40) were used as hosts when working with λgt11 clones. *E. coli* LE392 (20) was used as a host when working with λEMBL4 clones. The plasmids pUC18 and pUC19 were used in *E. coli* subclonings (39). *S. griseolus* ATCC 11796 and derivatives *S. griseolus* PH2001 and PH2042 (P. Harder, unpublished data) were the sources of cytochrome P-450$_{SU1}$, cytochrome P-450$_{SU2}$, and DNA encoding the *suaC* and *subC* genes.

Bacteriological methods. *E. coli* cultures were grown on LB agar and in LB or NZYCM broth at 30 to 42°C depending on the strain used (18). When used for selection, ampicillin was added at 100 μg/ml, tetracycline at 10 μg/ml, and kanamycin at 50 μg/ml. *E. coli* was transformed as described previously (6). *E. coli* strains lysogenic for λgt11 derivatives were made by infecting *E. coli* Y1089 at a multiplicity of infection of approximately 5 and screening surviving *E. coli* cells for sensitivity to 42°C (lysogens).

S. griseolus was grown in sporulation broth (0.1% [wt/vol] beef extract, 0.2% [wt/vol] tryptose, 1% [wt/vol] glucose, pH 7.1) at 30°C.

DNA isolation and manipulation. Total DNA from *S. griseolus* was isolated after lysozyme treatment in 150 mM NaCl–100 mM EDTA (pH 8.0)–10 mg of lysozyme per ml at 37°C for 2 h. Sodium dodecyl sulfate (SDS) was added to a concentration of 1%, and the mixture was heated at 65°C for 10 min to lyse the cells. CsCl (1.1 g/ml of lysate) was added along with approximately 100 μg of ethidium bromide per ml, and the solution was centrifuged in a Beckman VTi65 rotor for 12 h at 50,000 rpm. The DNA band was removed from the tube, and the ethidium bromide was extracted with NaCl · H_2O-saturated isopropanol and the CsCl was removed by repeated ethanol precipitations.

Plasmid DNAs were isolated either by rapid miniprep procedures (18) or by the alkaline lysis method and subsequent CsCl-ethidium bromide density gradient centrifugation (18). DNAs from λEMBL4 and λgt11 recombinants were isolated by the plate lysate method (18).

DNA manipulations, restriction digestions, and subcloning of DNAs were performed essentially as described previously (18). Nick translation of DNAs was performed with [α-^{32}P]dCTP as the label (29), and 5′ end labeling of oligonucleotides was performed as described previously (18).

DNA filter hybridizations. Southern transfers of DNA from agarose gels to nitrocellulose or Genescreen II (Dupont, NEN Research Products) were performed as described previously (33). DNA from bacteriophage plaques in petri dishes was affixed to nitrocellulose as described previously (18). When cloned DNA fragments were used as hybridization probes, hybridization was performed in 6× SSPE (1× SSPE is 0.18 M NaCl, 10 mM $NaPO_4$, 1 mM EDTA, pH 7.4)–5× Denhardt solution (7)–100 μg of boiled, sonicated calf thymus DNA per ml–50% (vol/vol) formamide at 42 to 52°C. Filters were washed four times for 5 min each in 2× SSPE–0.5% SDS at room temperature and then twice for 10 min each in 0.5× SSPE–0.5% SDS at 42°C. When an oligonucleotide probe was used, hybridization was performed in 5× SSPE–5× Denhardt solution–100 μg of boiled sonicated calf thymus DNA per ml–1% SDS at 50°C. Filters were washed four times for 5 min each in 6× SSPE–0.5% SDS at room temperature and twice for 10 min each in 6× SSPE–0.5% SDS at 60°C.

Construction and screening of bacteriophage λ DNA libraries. A λgt11 library of *S. griseolus* DNA was made as described previously (34). The library consisted of approximately 4.8 × 10^6 independent plaques (~30% with inserts). Polyclonal rabbit anti-cytochrome P-450_{SU1} antiserum (23) was preadsorbed to remove anti-*E. coli* antibodies by the pseudoscreening method (34). Screening of the λgt11-*S. griseolus* DNA library was performed as described previously, using ^{125}I-protein A to detect positive clones (34).

A λEMBL4 (10) library of 10- to 20-kilobase (kb) partially *Sau*3AI-digested DNA fragments of *S. griseolus* inserted into the *Bam*HI site of λEMBL4 was made (approximately 6.5 × 10^6 independent phage) (18). The library was screened with ^{32}P-labeled nick-translated DNA as a probe (18).

A size-fractionated library of *S. griseolus* 2.0-kb *Bam*HI DNA fragments was made in pUC19. *S. griseolus* PH2001 DNA (10 μg) was digested to completion with *Bam*HI and electrophoresed in a low-melting-point agarose gel. DNA from the 2.0-kb region of the gel was cut out and eluted. This DNA was then cloned into *Bam*HI-cut pUC19 DNA.

DNA sequencing. The 2.4-kb *Bam*HI DNA fragment containing the *suaC* gene was cloned in both orientations in pUC118, and a set of deletions (endpoints every ~0.1 kb) was made by exonuclease III digestion (14) of the plasmids. The 2.0-kb DNA fragment containing the *subC* gene was sequenced after subcloning it into M13. Single-strand DNA was made from each deletion plasmid or M13 clone and sequenced by the dideoxy-chain termination method (31). Sequencing reactions with 7-deaza-dGTP (19) and Sequenase (U.S. Biochemical Corp.) were used to give better reading of the high percent G+C *Streptomyces* DNA (19). The DNAs were sequenced on both strands, and the sequences were analyzed by using the University of Wisconsin DNA analysis programs (8).

Protein gel electrophoresis and Western blot analysis. *E. coli* lysogens were grown in 10 ml of LB broth at 30°C to an optical density of 0.5 at 600 nm, and then 50 μl of 0.1 M isopropyl-β-D-thiogalactopyranoside was added and the culture was incubated at 44°C for 10 min. The cultures were then shaken at 37°C for 1 h. The cells were pelleted at 5,000 × *g* for 5 min, washed with TBS (150 mM NaCl, 50 mM Tris chloride, pH 8.1), and then suspended in TBS plus 20% glycerol and broken by freezing and thawing twice. Cellular debris was removed by pelleting at 15,000 × *g* for 10 min. Proteins from plasmid-bearing (pCE320, pUC18-24a, or pUC19) strains of *E. coli* were prepared in the same manner as those from *E. coli* lambda lysogens except that cultures were grown at 37°C at all times before and after addition of isopropyl-β-D-thiogalactopyranoside. Protein concentrations were determined by the Bio-Rad protein assay with bovine serum albumin as a standard (4).

SDS gel electrophoresis of proteins was performed as described previously with 7.5% polyacrylamide gels (16). Proteins in these gels were electrophoretically transferred to nitrocellulose, and specific proteins were antigenically identified by the Western blot procedure (36) with alkaline phosphatase-conjugated anti-rabbit antiserum to identify the proteins to which the primary antibody bound.

TABLE 1. Comparison of the measured amino acid composition of cytochromes P-450_{SU1} and P-450_{SU2} with the amino acid composition predicted from the DNA sequences of *suaC* and *subC*[a]

Residue	No. of residues/mol			
	P-450_{SU1}	SuaC	P-450_{SU2}	SubC
Asx	40	37	38	36
Thr	28	30	31	33
Ser	18	19	14	14
Glx	48	44	39	35
Pro	28	27	29	33
Gly	33	29	30	27
Ala	48	43	52	45
Val	33	32	29	32
Met	7	7	8	9
Ile	13	18	14	17
Leu	48	48	50	49
Tyr	8	7	4	4
Phe	13	12	13	13
His	11	12	10	9
Lys	6	5	5	3
Trp[b]	2	2	2	3
Arg	34	30	49	41
Cys[c]	3	3	6	6

[a] Measured amino acid compositions of P-450_{SU1} and P-450_{SU2} are based on analysis of ~400 pmol of apoprotein, using a 24-h hydrolysis in 6 M HCl at 110°C, and the residues per mole calculation assumed a molecular weight of 45,000 for both proteins. The predicted residues per mole calculations for the SuaC and SubC proteins are from translation of the open reading frames for either protein as deduced from the DNA sequences of the *suaC* and *subC* genes (Fig. 5A and B).

[b] Tryptophan was estimated from the UV spectra of the apoproteins in 6 M guanidine hydrochloride by the method of Ragone et al. (27).

[c] Cysteine was measured as carboxymethylcysteine (9).

RESULTS

Characteristics of cytochromes P-450_{SU1} and P-450_{SU2}. Previously, it had been noted that the spectroscopic, chromatographic, and antigenic properties of cytochromes P-450_{SU1} and P-450_{SU2} were sufficiently different to suggest that the cytochromes were distinct gene products (23, 24). Purification of both cytochromes and analysis of the properties of the apoproteins prepared as in Fig. 1 allowed further comparison. Major differences in the proteins, particularly in the Tyr, Arg, and Cys content (Table 1), the N-terminal sequences (Fig. 2), and the readily resolvable chromatographic differences, under both native (24) and denaturing (Fig. 1) conditions confirmed that these proteins were distinct and not derived one from another by proteolysis, as has been

	10	20	30
$P450_{SU1}$	TDTAT TPQTT DAPAF PSNRS CPYQL XDGYA Q		
$P450_{SU2}$	TTAEX TAPPD ALTVP ASRAP GCPFD PAPDV TE		

FIG. 2. Amino-terminal sequencing of apocytochromes P-450. Sequences were based on the analysis of 544 pmol of $P\text{-}450_{SU1}$ and 620 pmol of $P\text{-}450_{SU2}$ (measured from amino acid composition). Based on these amounts of protein, sequencer amino acid yields on cycles 1 and 30 were 85 and 30%, respectively, for $P\text{-}450_{SU1}$ and 52 and 17%, respectively, for $P\text{-}450_{SU2}$.

shown to be the case with some *Streptomyces* P-450 cytochromes (37).

Cloning DNAs encoding cytochromes $P\text{-}450_{SU1}$ and $P\text{-}450_{SU2}$. A λgt11 library of *S. griseolus* was screened with polyclonal anti-$P\text{-}450_{SU1}$ antiserum. This antiserum has been shown to be specific for cytochrome $P\text{-}450_{SU1}$, with little cross-reactivity to cytochrome $P\text{-}450_{SU2}$ or other cytochromes P-450 tested (23) (also see Fig. 6). Six clones that reacted with the anti-$P\text{-}450_{SU1}$ antiserum were obtained. Preliminary restriction endonuclease mapping of the DNA

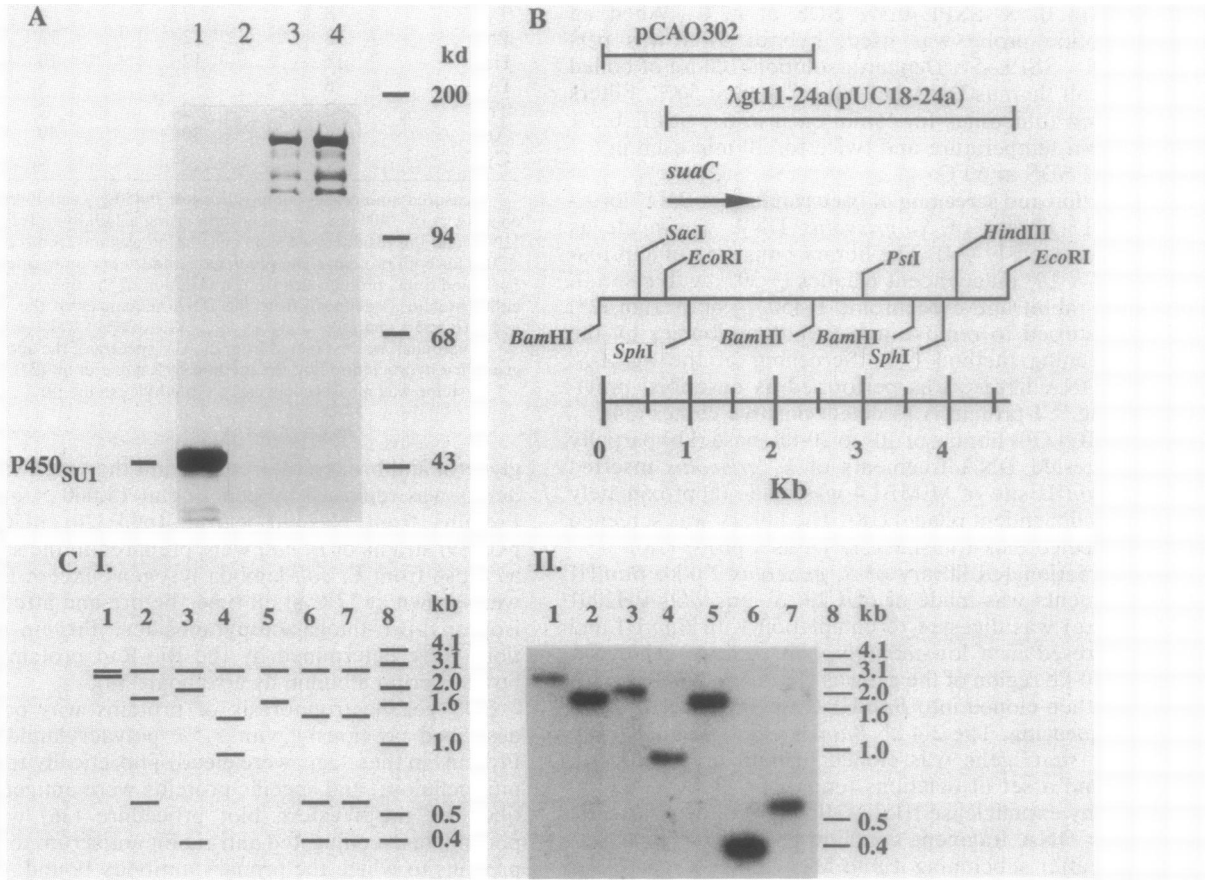

FIG. 3. (A) Western blot analysis of proteins from *E. coli* lysogenic for λgt11 or λgt11-24a with anti-$P\text{-}450_{SU1}$ antiserum. *E. coli* lysogens were grown up, induced, and lysed. Total protein (approximately 50 μg) was electrophoresed in an SDS-7.5% polyacrylamide gel, transferred to nitrocellulose, and probed with anti-$P\text{-}450_{SU1}$ antiserum. Lanes: 1, 1 μg of $P\text{-}450_{SU1}$ protein; 2, protein from an *E. coli* lysogen of λgt11; and 3 and 4, protein from two independent *E. coli* lysogens of λgt11-24a. The positions and sizes in kilodaltons (kd) of prestained marker proteins (Bethesda Research Laboratories) are indicated. (B) Restriction endonuclease map of the DNA of the *S. griseolus* chromosome surrounding the *suaC* gene. Also indicated are the DNA segments of this region in pCAO302 and λgt11-24a(pUC18-24a). (C) Probing of pCAO302 with a 16-fold-degenerate oligonucleotide derived from amino acids 6 to 12 of the amino-terminal amino acid sequence of purified $P\text{-}450_{SU1}$ (Fig. 2). The oligonucleotide used contained only G or C at position 3 of each codon because of the highly biased codon usage of *Streptomyces* species (3). The following oligonucleotide sequence was used:

```
5'-ACG/C-CCG/C-CAG-ACG/C-ACG/C-GAC-GC-3'
  Thr6 - Pro -Gln- Thr - Thr -Asp-Ala12
```

In panel I is a diagram of the restriction endonuclease fragments of the digests, and in panel II is the autoradiograph of the nitrocellulose filter upon which these DNAs were transferred and probed with the 5'-^{32}P-labeled oligonucleotide. All digestions are of pCAO302. Lanes: 1, *Bam*HI; 2, *Bam*HI plus *Eco*RI; 3, *Bam*HI plus *Sac*I; 4, *Bam*HI + *Sph*I; 5, *Bam*HI plus *Eco*RI plus *Sac*I; 6, *Bam*HI plus *Eco*RI plus *Sph*I; 7, *Bam*HI plus *Sac*I plus *Sph*I; 8, 1-kb ladder (Bethesda Research Laboratories) of marker DNAs. The sizes of the marker DNAs are indicated.

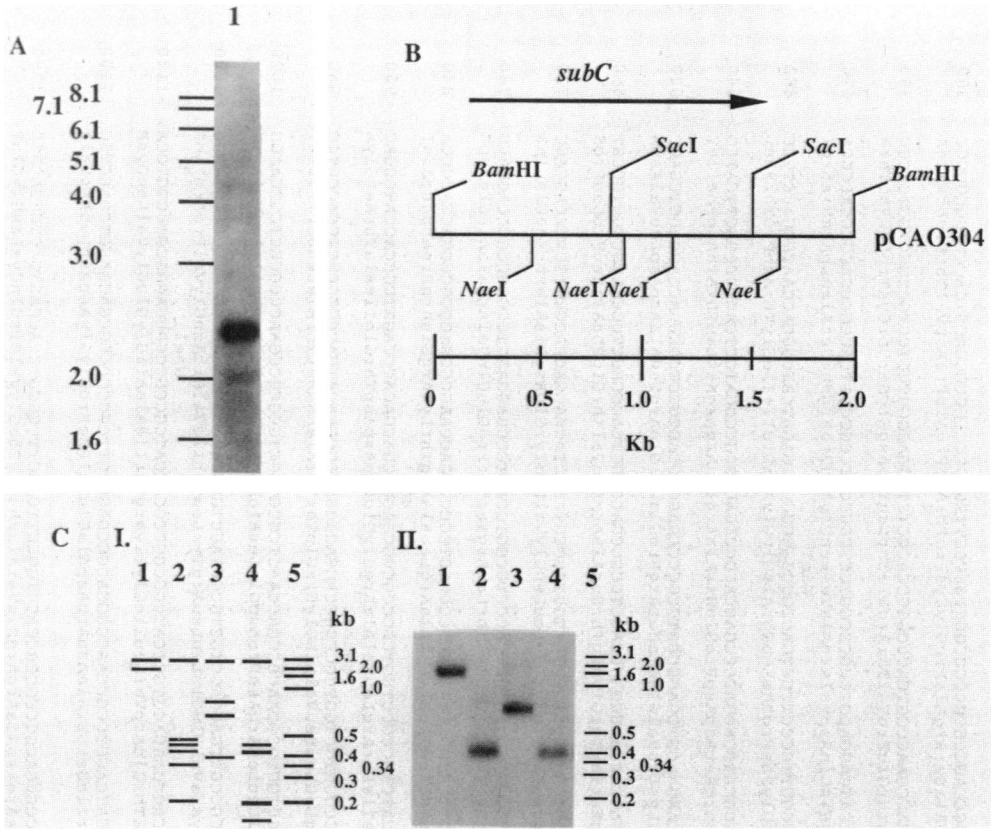

FIG. 4. Cloning of the *subC* gene. (A) Southern blot of *S. griseolus* DNA cut with *Bam*HI and probed with pCAO302. The sizes of the marker DNAs (in kilobases) are indicated. (B) Restriction endonuclease map of pCAO304 indicating the position of the *subC* gene coding sequence. (C) Restriction endonuclease digests of pCAO304 electrophoresed on a 1.5% agarose gel probed with a 16-fold-degenerate 26-mer oligonucleotide derived from amino acids 22 to 30 of the NH_2-terminal amino acid sequence of purified P-450_{SU2} (Fig. 2). The oligonucleotide used contained only G or C at position 3 of each codon because of the highly biased coding usage of *Streptomyces* species (2). The following oligonucleotide was used:

5′-TGC-CCG/C-TTC-GAC-CCG/C-GCG/C-CCG/C-GAC-GT-3′
Cys_{22}- Pro -Phe-Asp- Pro - Ala - Pro -Asp-Val_{30}

In panel I is a diagram of the restriction endonuclease fragments of the digest, and in panel II is the autoradiograph of the filter upon which these DNAs were transferred and probed with the 5′-^{32}P-labeled oligonucleotide as described in the Materials and Methods. All digestions are of pCAO304. Lanes: 1, *Bam*HI; 2, *Bam*HI plus *Nae*I; 3, *Bam*HI plus *Sac*I; 4, *Bam*HI plus *Nae*I plus *Sac*I; 5, 1-kb ladder (Bethesda Research Laboratories) of marker DNAs. The sizes of the marker DNAs are indicated.

inserts from these clones indicated that they were all from the same region of the *S. griseolus* chromosome (data not shown). We chose one of the antibody-positive phage, λgt11-24a, with which to make a lysogen of *E. coli* Y1089. Protein from the induced lysogen showed several, large 130- to 160-kDa protein bands that cross-reacted with anti-P-450_{SU1} antibody (Fig. 3A). This was consistent with it encoding a β-galactosidase fusion protein (β-galactosidase, 114 kDa; P-450_{SU1}, ~45 kDa) containing at least part of the P-450_{SU1}-coding sequence.

The insert from λgt11-24a (Fig. 3B) was subcloned into pUC18, creating pUC18-24a, and this plasmid was used to probe a λEMBL4 library of *S. griseolus* DNA to obtain additional clones from this region of the *S. griseolus* chromosome. A 2.4-kb *Bam*HI fragment was isolated from one of these λ clones that spanned the junction of the fusion protein encoding fragment from λgt11-24a, and it was subcloned into pUC18, creating pCAO302 (Fig. 3B). Restriction digests of the 2.4-kb *Bam*HI fragment from pCAO302 were probed with an eightfold-degenerate 20-mer oligonucleotide derived from the NH_2-terminal amino acid sequence of cytochrome P-450_{SU1} (Fig. 2). This showed that the NH_2 terminus of P-450_{SU1} was located within a 0.4-kb *Eco*RI-*Sph*I DNA restriction endonuclease fragment in pCAO302 (Fig. 3C). Since cytochrome P-450_{SU1} has a molecular weight of approximately 45,000, it should be encoded by an open reading frame encoding ~400 to 420 amino acids or be about 1,200 to 1,260 base pairs long. Assuming that the orientation of the gene shown in Fig. 3B was as determined from the λgt11-24a clone, the 2.4-kb *Bam*HI fragment in pCAO302 was expected to encode the complete cytochrome P-450_{SU1} protein.

When the 2.4-kb *Bam*HI DNA fragment containing the *suaC* gene was used to probe total *S. griseolus* DNA cleaved with *Bam*HI, a weak 2.0-kb band hybridized to the probe in addition to the expected strong 2.4-kb band (Fig. 4A). Although polyclonal antibodies to cytochromes P-450_{SU1} and P-450_{SU2} do not show strong cross-reactivity, we still suspected that the genes for the two proteins might show homology at the DNA level. In mammals, multiple cytochromes P-450 induced by the same compound often show significant homology at the DNA level (11, 35). Thus, we

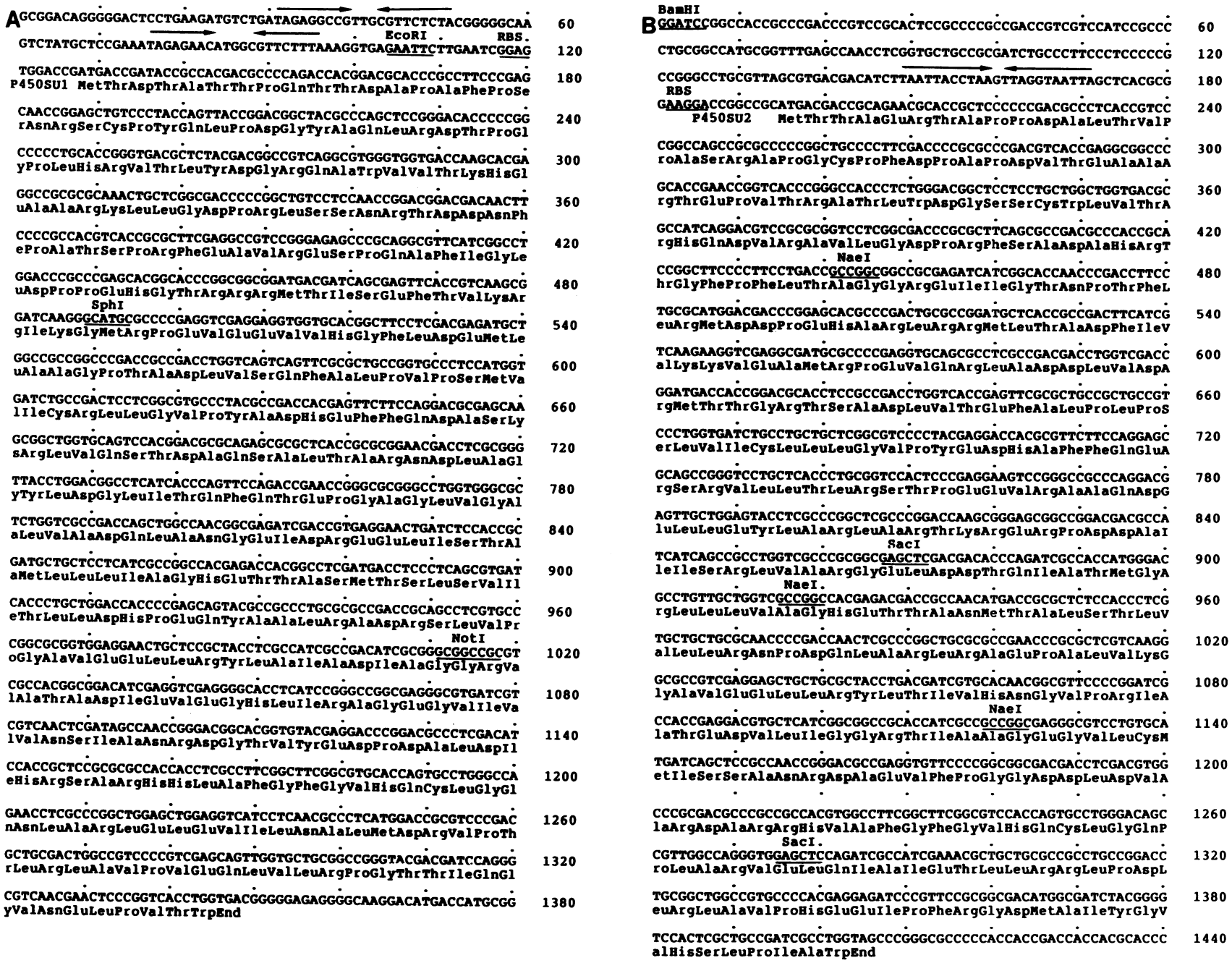
A
EcoRI
RBS.
P450SU1 MetThrAspThrAlaThrThrProGlnThrThrAspAlaProAlaPheProSe
SphI
NotI
B
BamHI
RBS
P450SU2 MetThrThrAlaGluArgThrAlaProProAspAlaLeuThrValP
NaeI
SacI
NaeI
SacI

FIG. 5. DNA sequences of the *suaC* (A) and *subC* (B) genes. The deduced amino acid sequences of cytochromes P-450$_{SU1}$ and P-450$_{SU2}$ are indicated along with the ribosome-binding sites (RBS) for each coding sequence. Inverted repeat sequences upstream of the *suaC* and *subC* coding sequences are indicated by arrows. The nucleotide sequence data reported here have been submitted to GenBank and assigned the accession numbers M32238 (*suaC*) and M32239 (*subC*).

cloned the DNA that cross-hybridized to the *suaC* gene. A library of 2.0-kb *S. griseolus* PH2001 *Bam*HI DNA fragments was made in pUC19 and was screened with the *suaC*-containing 4.0-kb *Eco*RI DNA fragment from λgt11-24a to identify clones containing the cross-hybridizing DNA. DNA from one such clone, pCAO304, was probed with an eightfold-degenerate 26-mer oligonucleotide derived from the NH_2-terminal amino acid sequence of cytochrome P-450$_{SU2}$ (Fig. 2 and 4C). Specific hybridization of the oligonucleotide to a 0.45-kb *Bam*HI-*Nae*I restriction endonuclease fragment indicated that the DNA contained at least the amino terminus of the *subC* gene. DNA sequence analysis below showed that this DNA fragment contained the complete coding sequence for cytochrome P-450$_{SU2}$.

DNA sequence of *suaC* (P-450$_{SU1}$) and *subC* (P-450$_{SU2}$) genes. The DNA sequence of the *suaC* and *subC* genes was determined (Fig. 5A and B). An open reading frame that began with the same 31 NH_2-terminal amino acids determined from purified cytochrome P-450$_{SU1}$ (excluding the initiation methionine) (Fig. 2) was found on pCAO302 in the region between the *Eco*RI and *Sph*I sites indicated by oligonucleotide probing. This open reading frame extended from nucleotides 128 to 1348 and encoded a protein of 44.3 kDa. DNA sequence analysis of the 2.0-kb insert in pCAO304 showed that it contained the *subC* gene encoding cytochrome P-450$_{SU2}$ (Fig. 5B). An open reading frame extended from nucleotides 195 to 1406 encoding a protein of 44.4 kDa that had the same 32 amino-terminal amino acids (excluding the initiation methionine) (Fig. 2) as purified cytochrome P-450$_{SU2}$. The overall amino acid composition of the encoded proteins was also very similar to that of purified cytochromes P-450$_{SU1}$ and P-450$_{SU2}$ (Table 1). The open reading frames for these two genes were typical of those from *Streptomyces* species in being 70.2% G+C for *suaC* and 72.4% for *subC* and having a very high percent G+C content in position 3 of the codons in the open reading frame (95.3 and 96.0%, respectively) (3). Both genes also had *Streptomyces* ribosome-binding sites 6 to 8 base pairs upstream of their initiation codons (2, 32).

We also used the open reading frames for *suaC* and *subC* to express cytochromes P-450$_{SU1}$ and P-450$_{SU2}$ in *E. coli* (Fig. 6). The plasmid pUC18-24a contained the *suaC* open reading frame directly downstream of the β-galactosidase promoter from pUC18. The plasmid pCE320 was created from pCAO304 by cleavage at its unique *Eco*NI site (base pairs 155 to 165 in Fig. 5B), 35 base pairs upstream of the *subC* start codon, and the unique *Hin*dIII site in the polylinker portion of the plasmid. The ends were filled in with the Klenow fragment of DNA polymerase, and the DNA was recircularized to form pCE320. The plasmid pCE320 contains the *subC* open reading frame directly downstream of the β-galactosidase promoter from pUC19. *E. coli* carrying pUC18-24a expressed a protein of the same size as cytochrome P-450$_{SU1}$ that was bound by anti-P-450$_{SU1}$ antiserum (Fig. 6A). Similarly, *E. coli* carrying pCE320 expressed

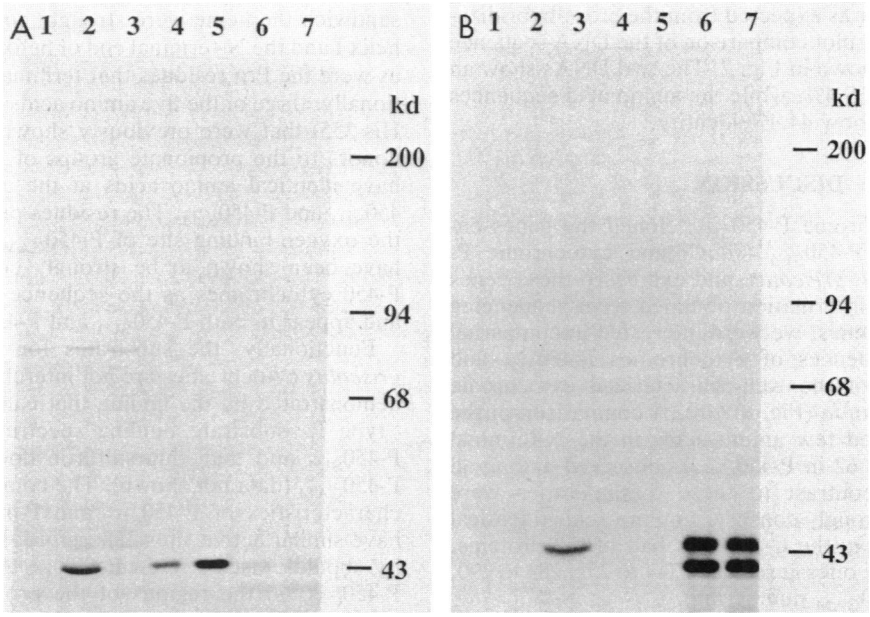

FIG. 6. Expression of the *suaC* and *subC* open reading frames in *E. coli*. Purified cytochromes P-450$_{SU1}$ or P-450$_{SU2}$ or total protein from lysed, plasmid-bearing strains of *E. coli* was run on SDS-polyacrylamide gels, transferred to nitrocellulose, and probed with antibody to either cytochrome P-450$_{SU1}$ (A) or P-450$_{SU2}$ (B). Lanes: 1, 100 μg of protein from *E. coli* containing pUC19; 2, 100 μg of protein from *E. coli* containing pUC18-24a; 3, 100 μg of protein from *E. coli* containing pCE320; 4, 100 ng of P-450$_{SU1}$; 5, 250 ng of P-450$_{SU1}$; 6, 100 ng of P-450$_{SU2}$; 7, 250 ng of P-450$_{SU2}$. The positions and sizes in kilodaltons (kd) of prestained marker proteins (Bethesda Research Laboratories) are indicated.

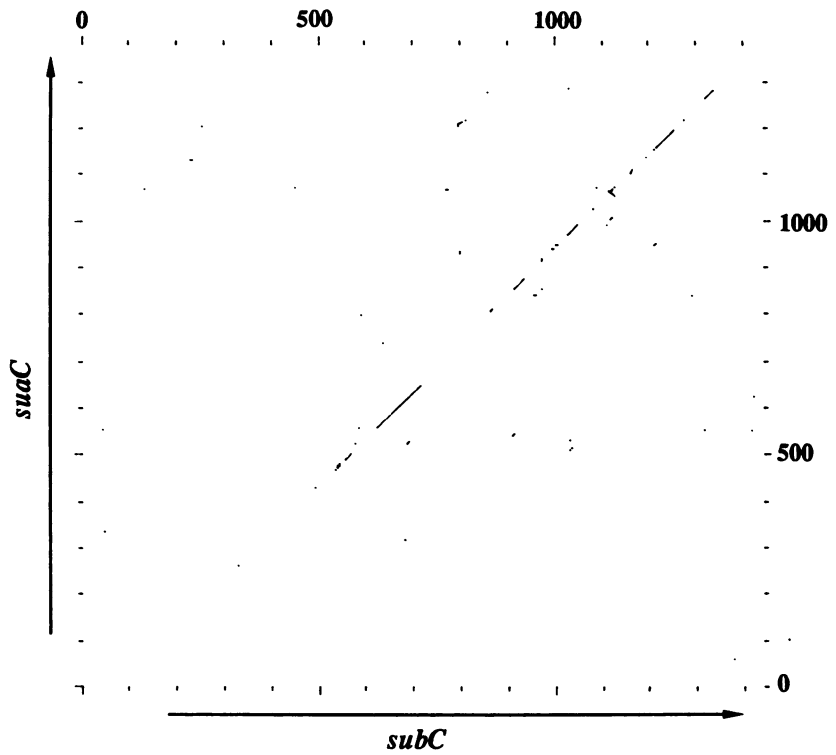

FIG. 7. Dot-plot comparison of the *suaC* and *subC* DNA sequences. The DNA sequences for *suaC* and *subC*, from Fig. 5A and B, were compared by using the program COMPARE with a window size of 21 and a stringency of 16 (8). the resulting points were plotted by using DOTPLOT (8). The *suaC* sequence is on the *y* axis, and the *subC* sequence is on the *x* axis. The protein-coding sequence of each gene are indicated by arrows.

a protein of the same size as cytochrome P-450_{SU2} that was bound by anti-P-450_{SU2} antiserum (Fig. 6B).

The DNA sequences of the *suaC* and *subC* genes are similar to one another as expected from the cross-hybridization of the two. A dot-plot comparison of the DNA sequence of the two genes is shown in Fig. 7. The two DNAs show an overall homology of 65.4%, while the amino acid sequences of the two proteins show 44.4% identity.

DISCUSSION

We purified cytochrome P-450_{SU2}, cloned the genes encoding cytochrome P-450_{SU1} (*suaC*) and cytochrome P-450_{SU2} (*subC*) from *S. griseolus*, and expressed these genes in *E. coli*. With the information obtained from sequencing the *suaC* and *subC* genes, we were interested in comparing the amino acid sequences of cytochromes P-450_{SU1} and P-450_{SU2} with that of the well-characterized cytochrome P-450_{CAM} from *P. putida* (Fig. 8) (38). A comparison of the three proteins showed few amino acids in the N-terminal region (residues 1 to 62 in P-450_{CAM}) conserved among all three proteins. In contrast to these dissimilarities were several patches of strongly conserved amino acids (identical in all three proteins) in the C-terminal half of the proteins, the major and notable ones at residues 244 to 252, 284 to 290, and 350 to 366 (P-450_{CAM} numbering).

The availability of a high-resolution (0.163-nm) crystal structure for cytochrome P-450_{CAM} (25, 26) enabled a higher-order structural analysis of the primary sequence comparison shown in Fig. 8. The relationship among the three proteins was strongest in the structural features involved in attaching and aligning the heme. The thiolate-proximal ligand of the heme (Cys-357 of P-450_{CAM}) and the residues forming a pocket around it (residues 350 to 360 in P-450_{CAM}) were strongly conserved; parts of two alpha helices that sandwich the heme were strongly conserved (the middle of helix I and the N-terminal end of helix L, indicated in Fig. 8), as were the Pro residues that terminate these helices. Additionally, three of the five amino acids (Arg-112, Arg-299, and His-355) that were previously shown to be hydrogen bond donors to the propionate groups of the heme in P-450_{CAM} have identical amino acids at the aligned positions of P-450_{SU1} and P-450_{SU2}. The residues proposed to form part of the oxygen-binding site of P-450_{CAM} (residues 248 to 252) have been shown to be strongly conserved in numerous P-450 cytochromes as the sequence -G/A-G-X-D/E-T- (26) and appear in both P-450_{SU1} and P-450_{SU2} as -A-G-H-E-T-.

Functionally, the substrates for P-450_{CAM} and the *S. griseolus* cytochromes are not interchangeable, most clearly demonstrated by the finding that camphor does not elicit a "type I" substrate binding spectrum from P-450_{SU1} and P-450_{SU2} and that chlorsulfuron does not elicit one from P-450_{CAM} (data not shown). The common substrate-product characteristics of P-450_{SU1} and P-450_{SU2} (23) imply they have similar active sites. Therefore, more sequence similarity should exist between P-450_{SU1} and P-450_{SU2} than with P-450_{CAM} in the regions of the protein that interact with substrate. Comparison of the three proteins revealed only one segment that fills both criteria of being in a region of substrate binding (from the P-450_{CAM} crystal structure) and being similar in both P-450_{SU1} and P-450_{SU2} (the positions analogous to 244 to 247 in helix I of P-450_{CAM}). Further analysis of the seven specific amino acids that are within 0.4

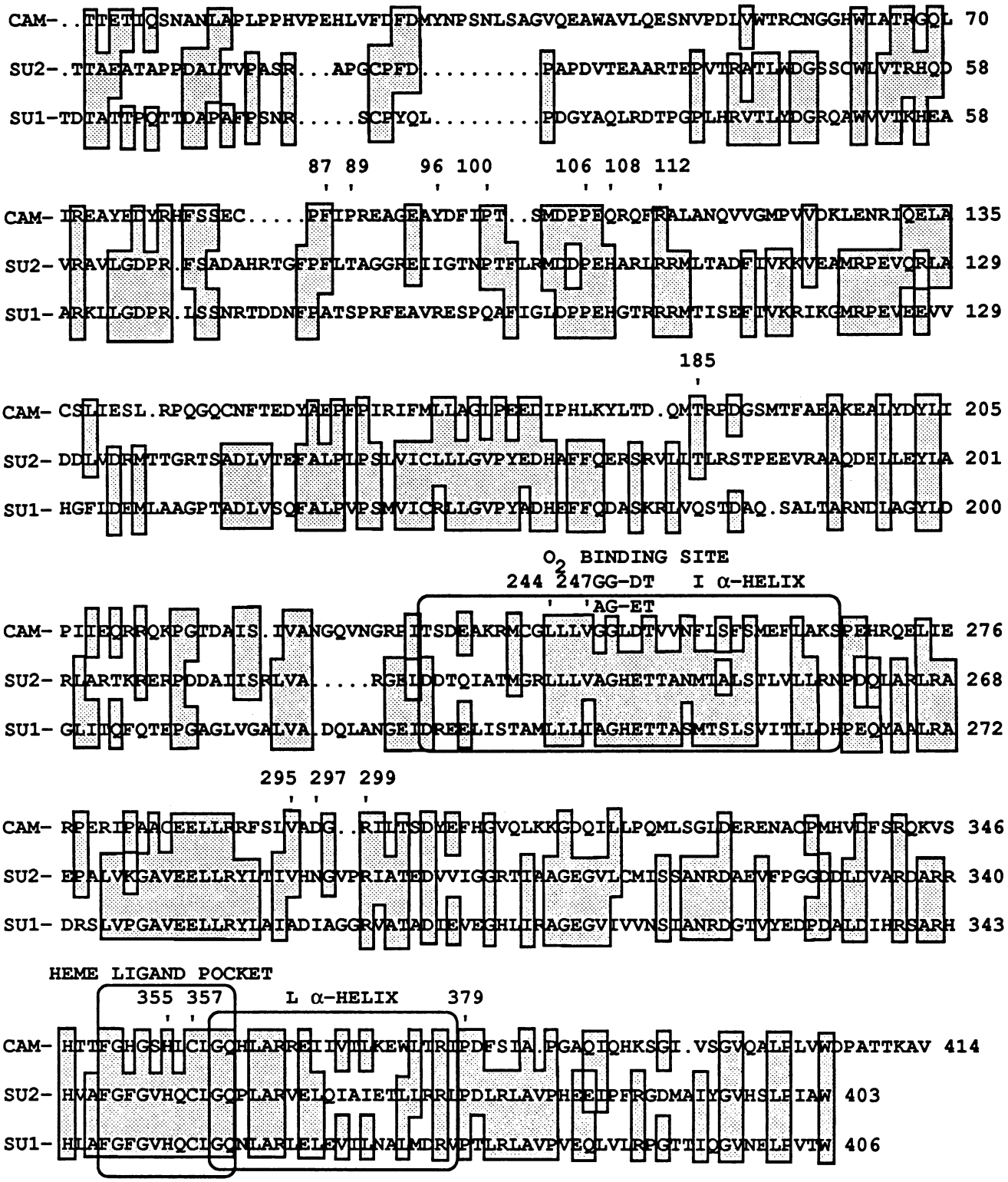

FIG. 8. Alignment of amino acid sequences of cytochromes P-450$_{CAM}$, P-450$_{SU1}$, and P-450$_{SU2}$. The amino acid sequence of P-450$_{CAM}$ (38) and those of P-450$_{SU1}$ and P-450$_{SU2}$ deduced from the DNA sequences of the *suaC* and *subC* genes were aligned by using the GAP program of the University of Wisconsin Genetics Computer Group Package (8). Pairwise comparisons were made with GAP by using a gap weight of 3.0 and a length weight of 0.1, and the gapped sequences were then compared with each other, again by using GAP. Final alignment of the sequences was done manually to remove redundant gaps. The positions of selected amino acids in the P-450$_{CAM}$ sequence are indicated above those amino acids: residues 87, 96, 185, 244, 247, 295, and 297 forming the substrate-binding site; three *cis*-prolines at 89, 100, and 106; the heme cysteine ligand (amino acid 357); and the hydrogen bond donors to the heme propionates (amino acids 108, 112, 297, 299, and 355). The amino acids composing helix I, helix L, and the heme ligand pocket of P-450$_{CAM}$ (amino acids 350 to 360) of P-450$_{CAM}$ and the corresponding aligned amino acids of P-450$_{SU1}$ and P-450$_{SU2}$ are boxed. The consensus amino acids (26) for the O_2-binding site are indicated at their corresponding positions (amino acids 248 to 252 of P-450$_{CAM}$). Shaded boxes indicate amino acids that are conserved between the aligned amino acid sequences of the P-450 cytochromes.

nm of the substrate in P-450$_{CAM}$ and of three *cis*-prolines important for forming part of the binding site revealed that five of the active site amino acids and two of the prolines are conserved in P-450$_{SU2}$. This similarity is notable given the overall lack of functional and sequence similarity between P-450$_{CAM}$ and P-450$_{SU2}$ and emphasizes the dissimilarity of P-450$_{SU1}$, which retains only one of the active site amino acids and one of the prolines.

In searching the National Biomedical Research Foundation protein data base, we found P-450$_{CAM}$ and P-450$_{BM-1}$ from *Bacillus megaterium* (13) to be the cytochromes P-450 most similar to P-450$_{SU1}$ and P-450$_{SU2}$. The overall amino acid sequences of P-450$_{SU1}$ and P-450$_{SU2}$ were shown to be, respectively, 29.7 and 29.4% identical to that of P-450$_{BM-1}$ and 25.9 and 27.8% identical to that of P-450$_{CAM}$. This, however, was considerably less similarity than had been found between members of cytochrome P-450 gene families in mammalian cells, in which members of a P-450 gene family have been defined as those having >36% amino acid identity (21). No P-450 amino acid sequence examined had >29.7% identity to that of P-450$_{SU1}$ and P-450$_{SU2}$, and thus these cytochromes P-450 are members of a new P-450 gene family by the criterion of Nebert et al. (21). They were shown to be members of the same P-450 gene family since the two proteins show an overall amino acid identity of 44.4%. Using the nomenclature of Nebert et al. (21), P-450$_{SU1}$ and P-450$_{SU2}$ were named P-450CVA1 and P-450CVB1, respectively.

Although the amino acid sequences of P-450$_{CAM}$ and the two herbicide-inducible cytochromes P-450 from *S. griseolus* showed regions of similarity to one another, a comparison of the DNA sequences encoding either of the two *Streptomyces* proteins with that coding for P-450$_{CAM}$ showed little similarity (data not shown). This was due, at least in part, to the higher percent G+C content of DNA from P-450$_{SU1}$ and P-450$_{SU2}$ of *Streptomyces* origin (70.2% G+C) compared with that of P-450$_{CAM}$ (38) from *P. putida* (59% G+C), resulting in a strong percent G+C bias in position 3 for each codon of the two *Streptomyces* genes. A similar lack of DNA sequence homology was seen between the DNA sequences of the genes encoding P-450$_{BM-1}$ (13) and *suaC* or *subC*. The DNA sequences encoding P-450$_{SU1}$ and P-450$_{SU2}$ did show significant homology as demonstrated by cross-hybridization of the two (Fig. 4A) and in the dot-plot comparison shown in Fig. 7. Thus, it is possible that the two *S. griseolus* cytochrome P-450 genes, *suaC* and *subC*, arose from one another by a gene duplication. Restriction endonuclease mapping of the two genes, however, did not show them to be near one another (>10 kb) on the *S. griseolus* genome (data not shown).

With the availability of the two genes encoding these herbicide-inducible cytochromes P-450, we intend to study how these genes are regulated by sulfonylureas as an example of a response of a microbe to chemicals foreign to its environment. Numerous sulfonylurea compounds of varied structures act as inducers of these genes in *S. griseolus*. We do not, however, know what is the natural inducer(s) of these genes. We also would like to understand the mechanism of the induction of *suaC* and *subC* since in wild-type *S. griseolus* the levels of cytochrome P-450$_{SU1}$ and P-450$_{SU2}$ proteins go from essentially undetectable before induction to greater than 1% of the soluble protein after induction with the sulfonylurea chlorimuron ethyl.

ACKNOWLEDGMENTS

We thank A. Stearrett, W. Alphin, and C. Maloney for technical assistance. We are also indebted to R. Kutny and M. Doleman for amino acid sequence and composition determinations of P-450$_{SU1}$ and P-450$_{SU2}$ proteins.

LITERATURE CITED

1. **Bergman, T., M. Carlquist, and H. Jornvall.** 1986. Amino acid analysis by high performance liquid chromatography of phenylthiocarbamyl derivatives, p. 45–53. *In* B. Wittmann-Liebold, J. Salnikow, and V. A. Erdmann (ed.), Advanced methods in protein microsequence analysis. Springer-Verlag, Berlin.
2. **Bibb, M. J., and S. N. Cohen.** 1982. Gene expression in *Streptomyces*: construction and application of promoter probe vectors in *Streptomyces lividans*. Mol. Gen. Genet. **187**:265–277.
3. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30**:157–166.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
5. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
6. **Cohen, S. N., A. C. Y. C. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.
7. **Denhardt, D. T.** 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **138**:179–207.
8. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
9. **Fontana, A., and E. Gross.** 1986. Fragmentation of polypeptides by chemical methods, p. 68–120. *In* A. Darbre (ed.), Practical protein chemistry—a handbook. John Wiley & Sons, Ltd., New York.
10. **Frischauf, A. M., H. Lehrach, A. Poustka, and N. J. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.
11. **Gonzalez, F. J., P. I. MacKenzie, S. Kimura, and D. W. Nebert.** 1984. Isolation and characterization of full-length mouse cDNAs and genomic clones of 3-methylcholanthrene inducible cytochrome P_1-450 and P_3-450. Gene **29**:281–292.
12. **Gunsalus, I. C., and G. C. Wagner.** 1978. Bacterial P450$_{CAM}$ methylene monooxygenase components: cytochrome M, putidaredoxin and putidaredoxin reductase. Methods Enzymol. **52**: 166–188.
13. **He, J.-S., R. T. Ruettinger, H.-M. Liu, and A. J. Fulco.** 1989. Molecular cloning, coding nucleotides and the deduced amino acid sequence of P450BM-1 from *Bacillus megaterium*. Biochem. Biophys. Acta **1009**:301–303.
14. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**: 351–359.
15. **Katagiri, M., B. N. Ganguli, and I. C. Gunsalus.** 1968. A soluble cytochrome P450 functional in methylene oxidation. J. Biol. Chem. **243**:3543–3546.
16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
17. **Machleidt, W., U. Borchart, and A. Ritonja.** 1986. Solid phase microsequencing: procedures and their potential for practical sequence analysis, p. 91–105. *In* B. Wittmann-Liebold, J. Salnikow, and V. A. Erdmann (ed.), Advanced methods in protein microsequence analysis. Springer-Verlag, Berlin.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mizusawa, S., S. Nishimura, and F. Seela.** 1986. Improvement of

the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.

20. **Murray, N. E., W. J. Brammer, and K. Murray.** 1977. Lambdoid phages that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. **150**:53–61.
21. **Nebert, D. W., M. Adesnik, M. J. Coon, R. W. Estabrook, F. J. Gonzalez, F. P. Guengerich, I. C. Gunsalus, E. F. Johnson, B. Kemper, W. Levin, I. R. Phillips, R. Sato, and M. R. Waterman.** 1987. The P450 gene superfamily: recommended nomenclature. DNA **6**:1–11.
22. **O'Keefe, D. P., and K. J. Leto.** 1989. Cytochrome P-450 from the mesocarp of avocado (*Persea americana*). Plant Physiol. **89**:1141–1149.
23. **O'Keefe, D. P., J. A. Romesser, and K. J. Leto.** 1987. Plant and bacterial cytochromes P450: involvement in herbicide metabolism. Recent Adv. Phytochem. **21**:151–173.
24. **O'Keefe, D. P., J. A. Romesser, and K. J. Leto.** 1988. Identification of constitutive and herbicide inducible cytochromes P450 in *Streptomyces griseolus*. Arch. Microbiol. **149**:406–412.
25. **Poulos, T. L., B. C. Finzel, I. C. Gunsalus, G. C. Wagner, and J. Kraut.** 1985. The 2.6 Å crystal structure of *Pseudomonas putida* cytochrome $P450_{CAM}$. J. Biol. Chem. **260**:16122–16130.
26. **Poulos, T. L., B. C. Finzel, and A. J. Howard.** 1987. High-resolution crystal structure of cytochrome $P450_{CAM}$. J. Mol. Biol. **195**:687–700.
27. **Ragone, R., G. Colonna, C. Balestrieri, L. Servillo, and G. Irace.** 1984. Determination of tyrosine exposure in proteins by second derivative spectroscopy. Biochemistry **23**:1871–1875.
28. **Rheinwald, J. G., A. M. Chakrabarty, and I. C. Gunsalus.** 1973. A transmissible plasmid controlling camphor oxidation in *Pseudomonas putida*. Proc. Natl. Acad. Sci. USA **70**:885–889.
29. **Rigby, P. W. J., M. Dieckman, C. Rhodes, and P. Berg.** 1977. Labelling of deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
30. **Romesser, J. A., and D. P. O'Keefe.** 1986. Induction of cytochrome P450 dependent sulfonylurea metabolism in *Streptomyces griseolus*. Biochem. Biophys. Res. Commun. **140**:650–659.
31. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
32. **Shine, J., and L. Dalgarno.** 1975. Determinant of cistron specificity in bacterial ribosomes. Nature (London) **254**:34–38.
33. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
34. **Snyder, M., and R. W. Davis.** 1986. Screening λgt11 expression libraries with antibody probes, p. 397–406. *In* T. Springer (ed.), Hybridoma technology in the biosciences and medicine. Plenum Publishing Corp., New York.
35. **Suwa, Y., Y. Mizukami, K. Sogawa, and Y. Fujii-Kuriyama.** 1985. Gene structure of a major form of phenobarbital-inducible cytochrome P-450 in rat liver. J. Biol. Chem. **260**:7980–7984.
36. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels in nitrocellulose sheets: procedures and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
37. **Trower, M. K., F. S. Sariaslani, and D. P. O'Keefe.** 1989. Purification and characterization of a soybean flour-induced cytochrome P-450 from *Streptomyces griseus*. J. Bacteriol. **171**:1781–1787.
38. **Unger, B. P., I. C. Gunsalus, and S. G. Sligar.** 1986. Nucleotide sequence of the *Pseudomonas putida* cytochrome $P450_{CAM}$ and its expression in *Escherichia coli*. J. Biol. Chem. **261**:1158–1163.
39. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene **33**:103–119.
40. **Young, R. A., and R. W. Davis.** 1983. Yeast RNA polymerase II genes: isolation with antibody probes. Science **222**:778–782.